Change in Total Energy of Comets Passing through the Solar System

EDGAR EVERHART

Physics Department, University of Connecticut, Storrs, Connecticut

(Received 27 August 1968)

Over 180,000 hypothetical long-period comets with orbits of random orientation are sent through a solar system having but one planet. Along each orbit is computed the change in the comet's total energy caused by the planet-comet interaction. Distributions of this quantity are presented wherein the comet's perihelion distances and inclinations are both specified. These changes for retrograde comets are found to be largest when their perihelion distance is near zero, but for direct comets there is a maximum effect when the comets' perihelion distances are comparable with the orbital radius of the planet. The one-planet case is solved in a dimensionless way so that simple scaling factors allow the solution to be applied to any number of planets in calculating individual orbits. However, the distribution in changes of energy with planets Mercury to Neptune included is shown to be practically the same as when Jupiter alone is considered. The distributions are symmetrical about zero but are clearly non-Gaussian in shape. At moderately large values of the energy change $E'$, the number of such changes falls off as $1/E'^{1/2}$. The predicted distribution agrees well with the values found for known long-period comets.

I. INTRODUCTION

GRAVITATIONAL attractions to the several planets alter the total energy of comets during their passage. There are many examples where a comet entering on a nearly parabolic orbit (total energy practically zero) has interacted with Jupiter with a gain in total energy for the comet so that its future orbit is hyperbolic. Such a comet cannot return, but other comets lose energy and their future orbits are elliptical.

Most previous calculations in this field treat in detail the changes in energy experienced by particular comets. For this purpose, one starts with the orbital elements in the vicinity of the sun (the osculating elements) and figures forward or backward in time to find the elements of the future or past orbit. The historic methods of Encke or Cowell may be used to trace the perturbed orbit in detail, and such detail is necessary for predicting the orbits and ephemerides of returning periodic comets. However, these methods are unnecessarily complicated if only the change in total energy is sought. From this quantity alone one knows whether the comet will return, and if so, at what intervals. Specialized solutions for this purpose have been given by Strömgren (1914), and these have been applied and developed by Makover (1955), Sekanina (1966a,b), Gallibina (1958), and others. Bilo and van de Hulst (1960) give an excellent review of this problem.

Studies using hypothetical comets can give over-all distributions of the expected energy changes. The problem is first solved for a solar system having a single planet and for comets whose initial orbits are parabolic, but taking account of all possible orientations of the comets' orbits, their perihelion distances, and times of passage. A start on the problem was made by van Woerkom (1948) who averaged a number of random orbits, and a more extensive study by Kerr (1961) formed distributions from about 14,000 orbits, sorting these according to perihelion distance. The present paper considers over 180,000 hypothetical comets with orbits randomly distributed. Distributions are formed wherein both the perihelion distance and the inclination are specified. Then the effects of all 8 planets, Mercury to Neptune, are considered and the expected distributions are derived for the actual solar system.

The present work is a continuation of two studies of comet distributions by the author: Everhart (1967a and 1967b), hereinafter to be called Papers I and II. These papers allow for observational selection and find intrinsic distributions of comets according to their perihelion distance, absolute magnitude, and orbital inclinations. To connect these distributions with various models for the origin of comets as, for example, those of Lyttleton (1953) and of Oort (1950), it is necessary to understand the statistical and cumulative effects of comet-planet interactions. Indeed, Lyttleton and Hammersley (1963) used the then available statistical information on these interactions in their study of the steady-state capture and loss rates among long-period comets. There is some relevance to the study of Kresák (1966) on the commensurability of comet periods with the period of Jupiter and to the studies of Opik (1963) on the lifetime of comets. However, these latter two studies are concerned with repeated interactions of periodic comets with Jupiter instead of single passages by many randomly distributed parabolic comets.

One may question whether observed long-period comets have random orbits. A nonuniform distribution with respect to argument of perihelion was accounted for in Paper II by observational selection effects, but some irregularities in the inclination distribution could not be so explained. Nonetheless, a random distribution of orbits is an appropriate and necessary assumption in the statistical treatment to be presented.

The present paper, Paper III of the sequence, uses computational procedures suitable to small and intermediate-sized perturbations, but which do not correctly apply to close encounters. A near collision with Jupiter that converts a long-period comet into a short-period comet requires different mathematical methods. These

1039
close encounters and their probabilities are treated separately in Paper IV, now in preparation.

II. PERTURBATION CALCULATIONS

Bilo and van de Hulst (1960) discuss the methods of calculating energy perturbations and derive the equations in careful detail. For almost all long-period comets it is sufficiently accurate to compute the changes in energy along an undisturbed parabolic path about the sun—even in those cases where the oscillating orbit is known to be either an ellipse of long period or a hyperbola with eccentricity near unity. Of course there is a large-energy perturbation when a comet passes near a planet, but there are also oscillatory contributions which persist as the comet is followed out to very large distances. These oscillations arise when the energy of the comet is referred to sun-centered coordinates, because the sun is itself in accelerated orbital motion about the center of gravity of the solar system, i.e., the barycenter. [An interesting plot of the actual motion of the sun with respect to the barycenter during the period 1950–1969 is given by Clemishaw (1968).] A shift of origin to the barycenter is made at an arbitrary large distance and this removes these oscillations.

A. Dimensionless Equations

An important advantage results when the equations are set in dimensionless form. Then it is only necessary to work out the solution for one planet and, by scaling according to mass and orbit size, sum the effects of any number of actual planets.

1. Dimensioned Scalars and Vectors:

- $a$, semimajor axis of comet's orbit;
- $a_p$, semimajor axis of planet's orbit;
- $q$, perihelion distance of comet;
- $t$, time;
- $G$, universal gravitation constant;
- $m_c$, mass of comet;
- $m_p$, mass of planet;
- $m_s$, mass of sun;
- $r$, comet's position vector relative to sun;
- $v$, comet's velocity vector relative to sun;
- $v_{ir}$, scalar magnitude of planet's velocity if in circular orbit;
- $u = \Delta(1/a)$, change in $1/a$ in reciprocal a.u.

2. Dimensionless Scalars and Vectors:

- $A = a/a_p$, comet's semimajor axis in units of planet's orbit radius;
- $Q = q/a_p$, comet's perihelion distance in units of planet's orbit radius;
- $T = v_{ir}/a_p$, dimensionless time—period of planet is $2\pi$ in this measure;
- $M = m_p/m_s$, mass of planet in units of sun's mass;
- $R = r/a_p$, comet's position vector, units of planet's orbit radius;
- $V = v/v_{ir}$, comet's velocity vector, units of planet's circular velocity;
- $R_p = r_p/a_p$, planet's position vector, magnitude unity for circular orbits;
- $V_p = v_p/v_{ir}$, planet's velocity vector, magnitude unity for circular orbits;
- $P$, unit vector from sun to perihelion position of comet;
- $S$, unit vector from sun to comet when comet has passed perihelion and has a true anomaly of $+90^\circ$ (usually called $Q$);

$$U = \Delta(1/A)/M$$

dimensionless change in comet's $1/A$ value per unit mass of the planet.

The actual total energy of a comet is $-Gm_c m_p/(2a)$, but the significant quantity is the change in energy per unit mass of the comet. The other factors are constant and it is sufficient to calculate the change in $1/a$. This is opposed in sign to the change in energy; when $u = \Delta(1/a)$ is positive the comet loses total energy. When measured in reciprocal astronomical units (a.u.$^{-1}$) the perturbations are ordinarily very small compared to unity and some authors quote values referred to the 6th decimal place in $u$.

To take advantage of the dimensionless equations and the symmetries of the problem, a special coordinate system is appropriate here: Termed "System I", the X and Y axes are chosen in the plane of the orbit of the perturbing planet with the fixed X axis pointing from the sun to the position the planet has at the moment when the comet is at perihelion. Thus at time $T = 0$ the planet is crossing the X axis and the comet is at perihelion.

The dimensionless equations for the comet's position and velocity are

$$\mathbf{R} = Q(1 - \eta) \mathbf{P} + 2Q \eta \mathbf{S} \quad (1)$$

and

$$\mathbf{V} = (2/Q)^{1/2}(\mathbf{S} - \eta \mathbf{P})/(1 + \eta^2), \quad (2)$$

where $\eta$ is found from Barker's equation

$$\eta + \eta^3/3 = T/(2Q^3). \quad (3)$$

In the case of a circular orbit, the planet's position and velocity are

$$\mathbf{R}_p = i \cos T + j \sin T \quad (4)$$

and

$$\mathbf{V}_p = -i \sin T + j \cos T,$$

with $i$ and $j$ being unit vectors in the X and Y directions.

B. The Calculation

The dimensionless perturbation $U$ to be found here is proportional to the change in total energy of the
comet during its entire orbit. To convert to this \( \alpha \) in \( \text{a.u.}^{-1} \) units, multiply \( U \) by the mass \( M \) of the planet in question (in units of the sun's mass) and divide by the mean radius of the planet's orbit (in a.u.). This total perturbation \( U \) is the sum of \( U_b \), the contribution received before perihelion, and \( U_u \), that received after perihelion. The perturbation \( U_u \) is found from three terms, and \( U_u \) is found similarly. Thus,

\[
U = U_u + U_b,
\]

\[
U_u = I_1 + I_2 + I_3,
\]

and

\[
U_b = I_1' + I_2' + I_3'.
\]

Here \( I_1 \) is the contribution when the comet is in the vicinity of both the sun and the planet, from perihelion to an arbitrary time \( T \), when the comet is well outside the perturbing planet. The term \( I_2 \) represents the change in the coordinate system 1. The axes of the system are shifted to the barycentric coordinates, that is, made at the time \( T = T_0 \). Finally, \( I_3 \) sums the barycentric contributions from \( T = T_0 \) to \( T = \infty \). The primed \( I \)'s refer to analogous terms before perihelion.

Let \( \rho = |R - R_p| \), let \( \lambda = |R| \), and let \( \psi = |V| \). Note further that \( |R_p| = 1 \) and \( |V_p| = 1 \). When the equations explained by Bilo and van de Hulst (1960) are recast in dimensionless form and specialized to the single planet case, then the result is

\[
I_1 = \int_0^{T_0} -2V \cdot [(R_p - R)\rho^2 - R_p]dT, \tag{6}
\]

\[
I_2 = 2\lambda^2[(R \cdot R_p\lambda^2 + 2V \cdot V_p\lambda^2 + 1], \tag{7}
\]

and

\[
I_3 = \int_0^{T_0} -2V \cdot [(R_p - R)\rho^2 - \lambda^2]dT + 3(R \cdot R_p\rho\lambda^2)]dT. \tag{8}
\]

In practice it is found that \( I_3 \) is unimportant provided that the dimensionless time \( T_0 \) is taken as large as \( 4\pi \), corresponding to two periods of the planet.

Both \( I_1 \) and \( I_2 \) depend on the choice made for \( T_0 \), but their sum, \( I_1 + I_2 \), is nearly independent of \( T_0 \) when \( T_0 \) is sufficiently large. The evaluation of \( I_1 \) requires care because the integrand oscillates through positive and negative values irregularly and, indeed, a close approach to the planet produces two near cusps as \( 4\pi \), corresponding to two periods of the planet. However, in cases where the comet approaches the planet closely, the size of the step is halved repeatedly. When the computer senses a fairly close approach the steps are taken up to 256 times smaller (or every 0.025 days) in considering Jupiter.

### Table I

<table>
<thead>
<tr>
<th>Comet</th>
<th>Name</th>
<th>Present calculation value ( 10^{-6} \text{a.u.}^{-1} )</th>
<th>Catalogue value ( 10^{-6} \text{a.u.}^{-1} )</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1937 IV</td>
<td>Whipple</td>
<td>+1248</td>
<td>+1308</td>
<td>a</td>
</tr>
<tr>
<td>1943 IV</td>
<td>Whipple-Bernasoni-Kalin</td>
<td>-820</td>
<td>-904</td>
<td>c</td>
</tr>
<tr>
<td>1944 IV</td>
<td>van Gent-Peltier-Daima</td>
<td>-501</td>
<td>-530</td>
<td>b</td>
</tr>
<tr>
<td>1946 I</td>
<td>Timmers</td>
<td>+438</td>
<td>+351</td>
<td></td>
</tr>
<tr>
<td>1946 VI</td>
<td>Jones</td>
<td>-56</td>
<td>-24</td>
<td>d</td>
</tr>
<tr>
<td>1948 I</td>
<td>Bester</td>
<td>-454</td>
<td>-391</td>
<td>e</td>
</tr>
<tr>
<td>1948 XI</td>
<td>Eclipse Comet</td>
<td>-816</td>
<td>-788</td>
<td></td>
</tr>
<tr>
<td>1949 I</td>
<td>Wirtanen</td>
<td>+40</td>
<td>+82</td>
<td></td>
</tr>
<tr>
<td>1949 IV</td>
<td>Baume-Bolk-Newkirt</td>
<td>+205</td>
<td>+219</td>
<td></td>
</tr>
<tr>
<td>1950 I</td>
<td>Johnson</td>
<td>+260</td>
<td>+291</td>
<td></td>
</tr>
<tr>
<td>1951 I</td>
<td>Minkowski</td>
<td>+219</td>
<td>+246</td>
<td></td>
</tr>
<tr>
<td>1953 III</td>
<td>Minkow-Honda</td>
<td>+250</td>
<td>+239</td>
<td></td>
</tr>
<tr>
<td>1958 III</td>
<td>Burnham</td>
<td>+1639</td>
<td>+1545</td>
<td></td>
</tr>
<tr>
<td>1959 III</td>
<td>Bester-Hoffmeister</td>
<td>-1620</td>
<td>-1342</td>
<td></td>
</tr>
<tr>
<td>1960 II</td>
<td>Burnham</td>
<td>-530</td>
<td>-455</td>
<td></td>
</tr>
</tbody>
</table>


The computation of the perturbation \( U_b \) before perihelion is carried out in a similar manner. It is necessary to change the sign of \( T_0 \) and \( \alpha \) in the limits of integration and to insert a minus sign in front of Eq. (7) to obtain the appropriate equations for the incoming branch of the orbit.

### C. Comparison with Previous Work

As a check on the procedure, the changes \( \Delta(1/\alpha) \) are computed for a number of comets where such a computation has already been reported.

For a given comet, and for each 8 planets in turn, it is first necessary to compute the vectors \( P \) and \( S \) in coordinate system 1. The axes of the system are differentially oriented for each planet and depend on the elements of the planet's orbit and its longitude in orbit at the time the comet is at perihelion. Circular orbits are taken for the planets and the dimensionless parameters established. Using Eqs. (5)-(8), \( U_b \) is found for each planet. The 8 contributions are added. Thus

\[
\Delta(1/\alpha) = \sum_{n=1}^{8} \frac{M_n}{a_{pm}} U_n, \tag{9}
\]

where \( M_n \) and \( a_{pm} \) are the mass and the mean orbital radius of each planet in turn. The calculation is done for the 15 most recent comets in Sekanina's catalogue (1966a) and the results are listed in Table I. The method outlined here may be among the simplest and best procedures for this computation, because the calculation for each planet uses a time scale and integration step appropriate to that planet. The calculation for 8 planets takes but 3 sec./comet on an IBM 360/65.
computer, so that it is unnecessary to neglect any planet.) The discrepancies seen in Table I between these values and those calculated elsewhere occur because here circular orbits are taken for the planets. It would be easy to replace Eqs. (4) by the correct equation for the planet's ellipse, but such refinement would not be in keeping with the present statistical study.

III. PERTURBATIONS CAUSED BY A SINGLE PLANET

Before adding together the effects of all planets it is necessary first to study the statistical effects of many random comets passing by one planet.

A. Symmetries

In coordinate system I the hypothetical comet has elements \( Q, \omega_0, \Omega_0, i_0 \), and the time of perihelion passage is zero. Here \( \omega_0, \Omega_0, \) and \( i_0 \) are defined as for the usual elements—argument of perihelion, longitude of ascending node, and inclination—except that the plane of reference is the planet's orbit plane rather than the ecliptic, and the point of reference is the position of the planet when the comet is at perihelion rather than the vernal equinox. The subscript "I" will be omitted henceforth, system I being understood.

It is evident from symmetry that for every hypothetical orbit [1] which has its perihelion point above the plane, there is a "mirror image" orbit [2] with its perihelion point below the plane. These orbits have the same changes in energy. That is, if \[ \omega_2 = \omega_1 + \pi, \] \( \Omega_2 = \Omega_1 + \pi, \) and \( i_2 = i_1 \), then

\[ U_2 = U_1. \tag{10} \]

Another symmetry is subtle and more restrictive: When the planet's orbit is circular, and there are two orbits such that

\[ \omega_2 = -\omega_1, \Omega_2 = -\Omega_1, \text{ and } i_2 = i_1, \]

then

\[ U_2 = -U_1. \tag{11} \]

For this second kind of symmetry the change \( U_1 \) occurring after perihelion in orbit [1] is the negative of the value \( U_2 \) occurring in orbit [2] before perihelion. (It turns out that at any particular instant \( -t \) before perihelion in case [1] all displacement and velocity vectors relating to the comet, planet, and sun are in the same relative orientation as they are at the instant \( +t \) after perihelion in [2], except that the comet's velocity vector is reversed. The consequence is that for each element of the comet's path in [1] the differential change in total energy is the negative of that for the symmetrical element in [2].) Kerr (1961) noted and made use of the symmetries of Eqs. (10) and (11) in his work.

The symmetry described in Eq. (11) has another consequence: Since for every orbit there is another orbit with an equal and opposite energy perturbation, the distribution of changes in energy for a large number of randomly oriented comets must be symmetrical and centered on zero. This is well known from other considerations—see Bilo and van de Hulst (1960), for example. Among actual comets, as tabulated by Brady (1965) and Sekanina (1966a), there is an excess of negative values of \( U \). Probably this is a result of selection, as Lyttleton and Hammersley (1963) have suggested. (See Sec. IVC for further discussion of this.)

In any case the theoretical average of zero for random comets is on a most sound footing. Note that here the discussion pertains to the total energy change including both the incoming and outgoing branches of the orbit. However the distribution of values \( U_1 \) along the (say) outgoing branch alone are neither symmetrical nor centered on zero, as treated in Sec. IIIIf below.

Anticipating the results of Paper IV, it is important to know that the symmetry result of Eq. (11) is of restricted validity and does not apply to the rare close encounters where there are large changes in energy and drastic changes in the comet's orbit. Indeed the \( U \) distribution is not symmetrical about zero if these rare cases are included.

B. Choice of Random Orbits

Two sets of distributions are to be found: First there are the cases wherein the perihelion distance \( Q \) is specified, but where the orbits are otherwise random.
Each orbit is specified by three different random numbers \(n_\omega, n_\Omega, n_\nu\), which are restricted to the range between \(-1\) and \(+1\). As is well known, random inclinations \(i\) have a distribution histogram that, for large populations, must vary as the sine of \(i\). It is not difficult to show that \(i = \arccos(n_\nu)\) generates such a sinusoidal population distribution of \(i\) values from the set of random numbers \(n_\nu\). The distribution in \(\omega\) and \(\Omega\) should be uniform, and the equations \(\omega = n_\omega \pi\) and \(\Omega = n_\Omega \pi\) generate uniformly random values of each over the range \(-\pi\) to \(+\pi\).

The calculation of the change \(U\) in total energy is carried out for an orbit chosen in such a way, and a new set of 3 random numbers selects the next orbit, and so on. For each value of \(Q\) there are 2000 such random orbits calculated. With 23 values of \(Q\), selected between 0.05 and 2.0, this requires 46,000 orbits.

The second set of cases is more extensive since one of 18 perihelion distances \(Q\) and one of 14 different inclinations \(i\) are specified for each distribution, a total of 252 cases. Here the \(\omega\) and \(\Omega\) values must be chosen to cover their ranges, preferably at random. About 550 orbits are computed for each \((Q,i)\) combination. A total of \(18 \times 14 \times 550\), or approximately 140,000 orbits, are studied in preparing this second set of distributions.

In preliminary trials the author, following Kerr (1961), tried covering the necessary ranges of \(\omega\) and \(\Omega\) by using a grid with the same size steps in each of these angles. However, this was found to cause large spurious steps in the resulting \(U\) distributions. The difficulty may be seen by an example: Suppose that the size of the grid steps is (say) \(4^\circ\) for both \(\omega\) and \(\Omega\), and that the calculation is carried out for particular values \((\omega_0,\Omega_0)\).

It turns out that the several other combinations \((\omega_0 + 4^\circ, \Omega_0 - 4^\circ), (\omega_0 - 4^\circ, \Omega_0 + 4^\circ), (\omega_0 + 8^\circ, \Omega_0 - 8^\circ), \ldots\) etc., all correspond to practically the same \(U\) values. The uniform grid does not pick a representative sample of orbits. The situation is very much better if the steps in \(\omega\) and \(\Omega\) are not too simply commensurate with each other. This improved procedure was actually followed, using steps of \(3.6^\circ\) in \(\omega\) and \(22.5^\circ\) in \(\Omega\). The computation of the 140,000 orbits was done first in point of time. From hindsight, that is from the later experience in computing the 46,000 random orbits of the other set, and the many thousands of random orbits of Paper IV, it would have been better if the 140,000 orbits had also been chosen at random \(\omega\) and \(\Omega\) values. However several trials showed the improvement to be marginal and the 140,000 calculations were not redone.

C. The Distribution in \(U\)

For each of the distributions \(h(U)\), a histogram may be formed. Here \(h(U) = \Delta N / \Delta U\), where \(\Delta N\) is the number of comets receiving an energy perturbation \(U\) in the range \(\Delta U\). Figure 1 shows such a histogram for \(Q=0.2\), one of the first set of cases where \(Q\) alone is specified but where the orbits are otherwise random. A logarithmic scale is chosen for \(h\) to better show the less frequent large perturbations. It is only necessary to plot the positive side since the distribution is symmetrical about zero. This histogram is normalized to have unit area to the right of zero, but contains the statistics of \(2000\) orbits. Under the scale of \(U\) on the abscissa are shown the values of \(n\) in a.u. \(^{-1}\) that would apply if the planet were Jupiter.

One of the 252 distributions of the second set of cases is plotted in Fig. 2, which shows the case where \(Q=0.1\) and \(i=171^\circ\). This is considerably broader than the case illustrated in Fig. 1; in fact, the distributions are always broader when they correspond to inclinations near \(0^\circ\) or near \(180^\circ\).

Among the distributions there is a wide diversity in shapes, but it is not practical to show 252 different figures. The most compact and useful way to summarize the results is to find an appropriate function which can be made to fit all cases, and then tabulate the parameters of that function.

Unfortunately the histograms do not fit a Gaussian or normal distribution because of the extensive tails found at large \(U\) values. Indeed the standard deviation, or root-mean-square (rms) value of \(U\), is sometimes 6 times larger than the median value of \(|U|\), but this factor should be 1.5 for a Gaussian distribution. Kerr (1961) tabulated rms values obtained by neglecting the tail completely (treating the distribution as Gaussian), whereas van Woerkom (1948) gave some average values that included the tail. It is not surprising that these papers, though each correct in what they tabulate, appear to disagree with each other.
D. Functional Fit to the Distribution in $U$

An appropriate function $h(U)$ to fit these distributions is

$$h(U) = G \exp[-(C/|U| - D)^2] + F/[U^3 + B^2]^{3/2}. \quad (12)$$

The first term is a Gaussian expression whose width depends mostly on $C$. The $D$ parameter shifts the center of the Gaussian and this helps accommodate a dip in the distribution often found near $U=0$. (See Fig. 2.) In most cases $D$ is set equal to unity. The asymmetry about $U=0$ is preserved because Eq. (12) has $|U|$ as the variable. This first term adjusts the fit at small $U$ values but quickly becomes negligible as $U$ increases.

The second term decreases slowly and fits the tail of the distribution. At large $U$ values (and within the framework of the present computation) it turns out that $h(U)$ must vary as $F/|U|^3$, and this dictates the asymptotic dependence. The inverse cube dependence on $U$ is derived in Paper IV, and there the parameter $F$ is evaluated analytically. [This analytic solution for $F$ applies strictly where $Q$ is less than about 0.8 and where both the $Q$ value and the inclination are specified. When $Q > 0.8$ then $F$ is regarded as an empirical parameter and is adjusted for best fit at each value of $Q$ and $i$. Having in hand a set of $F$ values for all $(Q,i)$ combinations it is easy to calculate a weighted average value of $F$ to use in those cases where only $Q$ is specified and where orbital inclination is random.]

The form of the second term is modified by the inclusion of a parameter $B$. The asymptotic dependence at large $U$ remains $F/|U|^3$, but an infinity at $U=0$ is prevented. In practice, $B$ provides a necessary adjustment at small and intermediate values of $U$.

Of the 4 empirical parameters ($G, C, D,$ and $B$) only 3 are independent because of the normalization requirement that $\int h(U) dU = 1$, integrated over all positive values of $U$. It may be shown that this requires that

$$G = 2C(1-F/B^2)J, \quad (13)$$

where erf is the complement of the error function, that is, erf($x$) = 1 - erf($x$).

The fitting procedure is explained in the Appendix. The solid lines in Figs. 1 and 2 show examples of the fit achieved using Eq. (12). The over-all shape is fitted well but not the small-scale irregularities. The adjustment is such that the logarithm of the histogram fits the logarithm of the curve as well as possible. Such a procedure gives as much importance to the few cases with large $U$ values as to the many cases with small $U$ values. An adjustment on a linear scale would have fitted better the many cases near $U=0$, but would have given no information near the tail of the curve.

Since the $F/|U|^3$ dependence which describes the tail of the curve is correct analytically, the results are valid to fairly large $U$ values. However, as discussed in Paper IV, the $1/|U|^3$ dependence breaks down for the rare near collisions. There is a value $U_m$ which is an approximate or order-of-magnitude upper limit on the perturbation calculations of the present paper. When $Q < 10$ these near collisions can occur and $U_m$ is tabu-

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|}
\hline
$q$ (a.u.) & $Q$ & $G$ & $C$ & $D$ & $F$ & $B$ & $U_d$ or $U_m$ \\
\hline
0.01 & 0.002 & 0.063 & 0.201 & 1.0 & 2.17 & 2.10 & 300 \\
0.03 & 0.005 & 0.161 & 0.254 & 1.0 & 2.19 & 2.49 & 300 \\
0.10 & 0.020 & 0.137 & 0.267 & 1.0 & 2.22 & 2.39 & 300 \\
0.26 & 0.050 & 0.111 & 0.277 & 1.0 & 2.27 & 2.57 & 300 \\
0.52 & 0.100 & 0.099 & 0.276 & 1.0 & 2.37 & 2.40 & 300 \\
1.04 & 0.200 & 0.093 & 0.290 & 1.0 & 2.61 & 2.30 & 300 \\
1.56 & 0.300 & 0.114 & 0.367 & 1.0 & 2.89 & 2.42 & 300 \\
2.08 & 0.400 & 0.177 & 0.480 & 1.0 & 3.23 & 2.86 & 300 \\
2.34 & 0.450 & 0.262 & 0.542 & 1.0 & 3.43 & 3.04 & 300 \\
2.60 & 0.500 & 0.306 & 0.563 & 1.0 & 3.67 & 5.70 & 300 \\
2.86 & 0.550 & 0.332 & 0.589 & 1.0 & 3.95 & 7.02 & 300 \\
3.12 & 0.600 & 0.359 & 0.633 & 1.0 & 4.30 & 6.43 & 200 \\
3.64 & 0.700 & 0.383 & 0.724 & 1.0 & 5.14 & 6.15 & 150 \\
4.16 & 0.800 & 0.406 & 0.757 & 1.0 & 5.65 & 5.45 & 100 \\
4.68 & 0.900 & 0.478 & 1.07 & 1.0 & 6.50 & 5.61 & 100 \\
5.20 & 1.000 & 0.513 & 1.07 & 1.0 & 9.20 & 6.44 & 40 \\
5.33 & 1.025 & 0.603 & 1.56 & 1.0 & 5.52 & 3.87 & 89 \\
5.46 & 1.050 & 0.625 & 1.54 & 1.0 & 4.68 & 3.72 & 62 \\
5.72 & 1.100 & 0.686 & 1.58 & 1.0 & 3.53 & 3.48 & 42 \\
6.50 & 1.250 & 0.000 & 1.04 & 1.0 & 1.04 & 1.02 & 16 \\
7.80 & 1.500 & 0.000 & 1.00 & 1.0 & 2.89 & 2.24 & 300 \\
10.4 & 2.000 & 0.000 & 0.14 & 0.17 & 0.14 & 0.17 & 300 \\
\hline
\end{tabular}
\caption{Coefficients $G$, $C$, $D$, $F$, and $B$ for use in Eq. (12) to compute the distribution $h(U)$. For this set the dimensionless perihelion distance $Q$ is specified but the orbits are otherwise random. The corresponding perihelion distance $q$ in a.u. considering Jupiter as the planet is also listed. The parameter $U_d$ is a rough limit beyond which Eq. (12) is not valid for $Q<1$. For $Q>1$ another parameter $U_m$ is tabulated beyond which the distribution is cut off and set equal to zero.}
\end{table}
### Table III. Coefficients for use in Eq. (12) to compute the distribution $k(U)$: The dimensionless perihelion distance $Q$ and the inclination $i$ (in degrees) are specified, but the orbits are random in the other orbital elements. For $Q > 1$ the table gives very approximate values $U_g$ beyond which Eq. (12) does not apply. For $Q > 1$ it lists a value $U_g$, at which the distribution is to be cut off and set equal to zero.

<table>
<thead>
<tr>
<th>$Q$ = 0.050, $g$ = 0.03 a.u.</th>
<th>$Q$ = 0.100, $g$ = 0.05 a.u.</th>
<th>$Q$ = 0.200, $g$ = 0.14 a.u.</th>
<th>$Q$ = 0.300, $g$ = 0.16 a.u.</th>
<th>$Q$ = 0.400, $g$ = 0.28 a.u.</th>
<th>$Q$ = 0.550, $g$ = 0.36 a.u.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i$</td>
<td>$G$</td>
<td>$C$</td>
<td>$D$</td>
<td>$F$</td>
<td>$B$</td>
</tr>
<tr>
<td>3</td>
<td>.015</td>
<td>.310</td>
<td>2.0</td>
<td>29.3</td>
<td>8.00</td>
</tr>
<tr>
<td>9</td>
<td>.184</td>
<td>.356</td>
<td>2.0</td>
<td>9.91</td>
<td>10.8</td>
</tr>
<tr>
<td>15</td>
<td>.209</td>
<td>.378</td>
<td>2.0</td>
<td>5.92</td>
<td>15.6</td>
</tr>
<tr>
<td>27</td>
<td>.220</td>
<td>.398</td>
<td>1.9</td>
<td>3.38</td>
<td>11.6</td>
</tr>
<tr>
<td>45</td>
<td>.233</td>
<td>.418</td>
<td>1.8</td>
<td>2.13</td>
<td>11.1</td>
</tr>
<tr>
<td>63</td>
<td>.169</td>
<td>.367</td>
<td>1.5</td>
<td>1.64</td>
<td>2.63</td>
</tr>
<tr>
<td>81</td>
<td>.174</td>
<td>.351</td>
<td>1.2</td>
<td>1.44</td>
<td>3.00</td>
</tr>
<tr>
<td>99</td>
<td>.132</td>
<td>.310</td>
<td>1.1</td>
<td>1.40</td>
<td>2.20</td>
</tr>
<tr>
<td>117</td>
<td>.111</td>
<td>.263</td>
<td>1.0</td>
<td>1.50</td>
<td>2.20</td>
</tr>
<tr>
<td>135</td>
<td>.102</td>
<td>.228</td>
<td>1.0</td>
<td>1.84</td>
<td>2.60</td>
</tr>
<tr>
<td>153</td>
<td>.083</td>
<td>.192</td>
<td>1.0</td>
<td>2.80</td>
<td>3.10</td>
</tr>
<tr>
<td>165</td>
<td>.062</td>
<td>.161</td>
<td>1.0</td>
<td>4.00</td>
<td>3.60</td>
</tr>
<tr>
<td>171</td>
<td>.050</td>
<td>.143</td>
<td>1.0</td>
<td>7.99</td>
<td>4.30</td>
</tr>
<tr>
<td>177</td>
<td>.041</td>
<td>.099</td>
<td>1.0</td>
<td>22.3</td>
<td>5.40</td>
</tr>
</tbody>
</table>

The dimensionless perihelion distance $Q$ and the inclination $i$ (in degrees) are specified, but the orbits are random in the other orbital elements. For $Q > 1$ the table gives very approximate values $U_g$ beyond which Eq. (12) does not apply. For $Q > 1$ it lists a value $U_g$, at which the distribution is to be cut off and set equal to zero.
<table>
<thead>
<tr>
<th>Q = 0.700</th>
<th>Q = 0.800</th>
<th>Q = 0.900</th>
<th>Q = 0.950</th>
<th>Q = 1.000</th>
<th>Q = 1.050</th>
<th>Q = 1.100</th>
<th>Q = 1.250</th>
<th>Q = 1.500</th>
</tr>
</thead>
<tbody>
<tr>
<td>( G )</td>
<td>( D )</td>
<td>( F )</td>
<td>( B )</td>
<td>( U_a )</td>
<td>( G )</td>
<td>( D )</td>
<td>( F )</td>
<td>( B )</td>
</tr>
<tr>
<td>3</td>
<td>.165</td>
<td>.450</td>
<td>1.0</td>
<td>149.3</td>
<td>19.3</td>
<td>300</td>
<td>3</td>
<td>.197</td>
</tr>
<tr>
<td>9</td>
<td>.170</td>
<td>.402</td>
<td>1.0</td>
<td>58.9</td>
<td>12.7</td>
<td>300</td>
<td>9</td>
<td>.196</td>
</tr>
<tr>
<td>15</td>
<td>.203</td>
<td>.408</td>
<td>1.0</td>
<td>29.5</td>
<td>12.5</td>
<td>300</td>
<td>15</td>
<td>.209</td>
</tr>
<tr>
<td>25</td>
<td>.251</td>
<td>.466</td>
<td>1.0</td>
<td>15.4</td>
<td>11.3</td>
<td>300</td>
<td>25</td>
<td>.262</td>
</tr>
<tr>
<td>45</td>
<td>.292</td>
<td>.589</td>
<td>1.0</td>
<td>8.00</td>
<td>8.33</td>
<td>300</td>
<td>45</td>
<td>.326</td>
</tr>
<tr>
<td>81</td>
<td>.417</td>
<td>.789</td>
<td>1.0</td>
<td>3.07</td>
<td>4.73</td>
<td>200</td>
<td>81</td>
<td>.412</td>
</tr>
<tr>
<td>99</td>
<td>.557</td>
<td>1.07</td>
<td>1.0</td>
<td>2.12</td>
<td>3.48</td>
<td>200</td>
<td>99</td>
<td>.526</td>
</tr>
<tr>
<td>117</td>
<td>.692</td>
<td>1.45</td>
<td>1.0</td>
<td>1.54</td>
<td>2.71</td>
<td>200</td>
<td>117</td>
<td>.292</td>
</tr>
<tr>
<td>153</td>
<td>.639</td>
<td>1.67</td>
<td>1.0</td>
<td>1.20</td>
<td>1.76</td>
<td>150</td>
<td>153</td>
<td>.292</td>
</tr>
<tr>
<td>171</td>
<td>.086</td>
<td>.543</td>
<td>1.0</td>
<td>1.16</td>
<td>1.71</td>
<td>80</td>
<td>171</td>
<td>.111</td>
</tr>
<tr>
<td>177</td>
<td>.043</td>
<td>.343</td>
<td>1.0</td>
<td>6.04</td>
<td>2.75</td>
<td>60</td>
<td>177</td>
<td>.043</td>
</tr>
</tbody>
</table>

\( q = 3.64 \text{ a.u.} \)

\( q = 4.16 \text{ a.u.} \)

\( q = 4.68 \text{ a.u.} \)

\( q = 4.94 \text{ a.u.} \)

\( q = 5.20 \text{ a.u.} \)

\( q = 5.46 \text{ a.u.} \)

\( q = 5.72 \text{ a.u.} \)

\( q = 6.50 \text{ a.u.} \)
TABLE III (continued)

<table>
<thead>
<tr>
<th>$Q=1.500$, $q=7.80$ a.u.</th>
<th>$Q=2.000$, $q=10.4$ a.u.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i$</td>
<td>$G$</td>
</tr>
<tr>
<td>3</td>
<td>0.0</td>
</tr>
<tr>
<td>9</td>
<td>0.0</td>
</tr>
<tr>
<td>15</td>
<td>0.0</td>
</tr>
<tr>
<td>27</td>
<td>0.0</td>
</tr>
<tr>
<td>45</td>
<td>0.0</td>
</tr>
<tr>
<td>63</td>
<td>0.0</td>
</tr>
<tr>
<td>81</td>
<td>0.0</td>
</tr>
<tr>
<td>99</td>
<td>0.0</td>
</tr>
<tr>
<td>117</td>
<td>0.0</td>
</tr>
<tr>
<td>135</td>
<td>0.0</td>
</tr>
<tr>
<td>153</td>
<td>0.0</td>
</tr>
<tr>
<td>165</td>
<td>0.0</td>
</tr>
<tr>
<td>171</td>
<td>0.0</td>
</tr>
<tr>
<td>177</td>
<td>0.0</td>
</tr>
</tbody>
</table>

When $Q>1.0$ such near collisions are impossible but another consideration appears: There is then found to be a sharp cutoff of $h(U)$ at a particular value $U_m$. That is

$$h(U)=0 \text{ for } U>U_m \text{ when } Q>1.$$  \hspace{1cm} (14)

Table II gives values of $G$, $C$, $D$, $B$, and $U_B$ or $U_m$ for all the sets where $Q$ alone is specified. In a compact form this summarizes the solution of the one-planet case.

The 252 distributions where both $Q$ and $i$ are specified are fitted by Eq. (12) with the parameters $G$, $C$, $D$, $B$, and $U_B$ or $U_m$ in Table III. This table provides a detailed solution of the one-planet case. Comments and precautions in the use of Tables II and III are given in the Appendix.

E. Percentiles

The $U$ values may be sorted and ranked according to percentiles and Fig. 3 shows the result for direct comets with inclinations of $3^\circ$. Here positive $U$ values are plotted versus $Q$ with contours at various percentiles. Thus the $1\%$ contour at $Q=0.3$ is seen to pass through $U=40$; i.e., at this $Q$ value and inclination $1\%$ of all comets (with random $\omega$ and $\Omega$) receive a positive $U$ value of 40 or more. Very large positive perturbations (where comets lose total energy) are common near $Q=1$, because it is possible there for the comet to accompany the planet for a sizeable portion of path. Of course the methods of this paper break down completely at such large $U$ values and the curves, shown dashed, are actually derived from the results of Paper IV. In that same dashed region there are comparatively few large negative $U$ values. However, wherever the contour lines are drawn solid, Fig. 3 would be the same if drawn for negative $U$ values.

Figure 4 shows a similar plot for inclinations of $90^\circ$ and the contours are everywhere lower. They are still lower in Fig. 5, which is for retrograde comets with...
1.2

Fig. 6. As in Fig. 3, except this set of $+U$ percentiles versus $Q$ is an over-all summary where the orbits have entirely random orientation in all 3 orbital angles.

An over-all view of the percentiles can be seen in Fig. 6 for which only the $Q$ value is specified and the orbits are otherwise random. This depicts the results of Table II. For example, the figure shows that 1% of all comets with $Q$ values of 0.2 lose energy corresponding to a $U$ value of $+10$ or more. There are two maxima seen—a weak one at $Q=0$ and a pronounced one at $Q=1$. The upturn near $Q=0$ is explained in Sec. IVD.

F. The Distribution in $U_b$ and $U_a$

Fabry (1894), Bilo and van de Hulst (1960), and Sekanina (1966a) discuss the average value of the energy changes $U_b$ along the before-perihelion branch and $U_a$ along the after-perihelion branch, considered separately. For a single planet of mass $m_p$ and orbit radius $a_p$ (and where $q < a_p$), they show that $\langle U_b \rangle = -2m_p/a_p$ and $\langle U_a \rangle = +2m_p/a_p$. The numerical value is $368 \times 10^{-6}$ a.u.$^{-1}$ when Jupiter alone is considered. (Adding the effects of the other planets can increase this average to $449 \times 10^{-6}$ a.u.$^{-1}$ as discussed in Sec. IVC below.) In terms of the dimensionless quantities, the above single-planet value is equivalent to the prediction that $U_b$ should be centered on $-2$ and $U_a$ centered on $+2$, provided that $Q$ is less than unity. That is, on the average, comets gain total energy before perihelion and lose total energy after perihelion. The example shown in Fig. 7 is for 2000 randomly oriented orbits that also have their $Q$ values distributed. (The distribution of $Q$ values corresponds to the $q$ distribution of Table IV as applied to Jupiter.) The histogram of Fig. 7 is centered on $+2$. This histogram agrees with the predicted center value, gives a measure of the width of the distribution and shows that the distribution has a most unsymmetrical skewed shape. For large values of $+U_a$ or $-U_a$ it goes as $1/|U_a|^3$. Further use of these results may be found in Sec. IVC.

IV. EFFECTS OF EIGHT PLANETS

A. Superposition

The effects of the several planets are additive as in Eq. (9), where it is seen that the $U$ values for each planet are proportional to the ratio of its mass to its mean orbital radius. For Jupiter this ratio is $1.84 \times 10^{-4}$. That is,

$$u = 1.84 \times 10^{-4} U$$

(15)

Fig. 7. Number distribution of 2000 $U_a$ values, the dimensionless changes in total energy after perihelion for comets in a solar system with one planet. The orbits of the comets have entirely random orientation and their $Q$ distribution is specified in the text. The before-perihelion values $U_b$ would be the same, statistically, except for a minus sign.

Table IV. The $q$ distribution of the 70 comets in Sekanina's catalogue (1966a) for which $u$ values have been computed. Steps in $q$ are 0.2 a.u. wide, centered on the value given. Where a step is missing the number $N$ in that step is zero.

<table>
<thead>
<tr>
<th>$q$</th>
<th>$N$</th>
<th>$q$</th>
<th>$N$</th>
<th>$q$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1 a.u.</td>
<td>2</td>
<td>1.1 a.u.</td>
<td>14</td>
<td>2.1 a.u.</td>
<td>3</td>
</tr>
<tr>
<td>0.3</td>
<td>5</td>
<td>1.3</td>
<td>3</td>
<td>2.3</td>
<td>5</td>
</tr>
<tr>
<td>0.5</td>
<td>1.5</td>
<td>2.5</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td>7</td>
<td>1.7</td>
<td>5</td>
<td>2.7</td>
<td>1</td>
</tr>
<tr>
<td>0.9</td>
<td>5</td>
<td>1.9</td>
<td>5</td>
<td>3.3</td>
<td>1</td>
</tr>
</tbody>
</table>

449 X $10^{-6}$ a.u.$^{-1}$ as discussed in Sec. IVC below.) In terms of the dimensionless quantities, the above single-planet value is equivalent to the prediction that $U_b$ should be centered on $-2$ and $U_a$ centered on $+2$, provided that $Q$ is less than unity. That is, on the average, comets gain total energy before perihelion and lose total energy after perihelion. The example shown in Fig. 7 is for 2000 randomly oriented orbits that also have their $Q$ values distributed. (The distribution of $Q$ values corresponds to the $q$ distribution of Table IV as applied to Jupiter.) The histogram of Fig. 7 is centered on $+2$. This histogram agrees with the predicted center value, gives a measure of the width of the distribution and shows that the distribution has a most unsymmetrical skewed shape. For large values of $+U_a$ or $-U_a$ it goes as $1/|U_a|^3$. Further use of these results may be found in Sec. IVC.
CHANGE IN ENERGY OF COMETS

The histogram drawn with a solid line is computed for 2000 random comets entering the solar system as described. The dotted curve is an empirical function whose equation and parameters are given in the text. The histogram traced by the dashed line is for 70 observed comets.

In units of \( a.u.^{-3} \) for this planet. The relative effects of the other planets compared to Jupiter are listed: Jupiter 1.0000; Saturn 0.1631; Venus 0.0185; Earth 0.0166; Uranus 0.0124; Neptune 0.0095; Mercury 0.0023; and Mars 0.0012.

Of course a given actual orbit has different \( U \) values for each planet, but nonetheless the above sequence of values is an over-all measure of the effect of each planet. In the calculation of a particular single orbit, as in the study of an actual comet, it is always necessary to include both Jupiter and Saturn. To get 1% accuracy it is clearly necessary to include at least 6 planets.

The situation is somewhat different if one desires to know the distribution of perturbations. The \( U_a \) and \( U_b \) distributions (after and before perihelion considered separately) are not centered on zero. There is a cumulative effect on the center value in adding on the several planets. A discussion of this complication is postponed to Sec. IVC below. However the \( U \) distribution, where \( U = U_a + U_b \), can be handled much more simply because the center value is zero. Adding the effects of the several planets therefore does not shift the distribution but only changes its width. It is plausible that the effect of the planets on the distribution width depends on the square root of the sum of the squares of these ratios. If the width for Jupiter alone is unity, then the distribution width for Jupiter and Saturn is \( \sqrt{(1.0)^2 + (0.16)^2} \), or 1.013, which is a negligible increase.

This tentative conclusion is supported by actual calculations: Using Monte Carlo methods, many \( U \) distributions are calculated, superimposing the effects of varying numbers of planets. Each distribution is formed by taking 1000 random orbits at specified \( q \) and \( i \) values. The effects of including Saturn and the rest of the planets is found to be completely negligible compared to the statistical fluctuations in the Monte Carlo treatment of Jupiter alone.

The result is that Tables II and III apply to the distributions for the entire Solar System. It is only necessary to scale the \( q \) values according to \( q = 5.2Q \) and the \( U \) values computed from Eq. (12) according to \( a = 1.84 \times 10^{-6}u \).

B. Comparison with Data: \( u \) Distribution

The over-all \( u \) distribution (all planets, all orientations all \( q \) values) is next calculated. For this purpose, another 2000 orbits are considered. The requirement of all orientations is met by using for each orbit three random numbers to pick the orbital angles, as described in Sec. III B. A fourth random number chooses the perihelion distance. The result depends, to some extent, on the distribution assumed for the \( q \) values, and here, rather arbitrarily, is used the same \( q \) distribution as that of the comets in Sekanina's catalogue (1966a). This is given in Table IV. The resulting 2000 \( u \) values form the histogram drawn with a solid line in Fig. 8. The value
of $F$ is computed as the appropriate weighted average of the $F$ values of Table II, and the other parameters of Eq. (12) are fitted to this histogram. It is found that $G = 0.093; C = 0.325; D = 1; F = 2.97$; and $B = 2.37$. The resulting distribution in $u$ (using $u = 1.84 \times 10^{-6} U$) is plotted as the dotted curve in Fig. 8. Both the histogram and the curve are over-all results suitable for comparison with the available information on long-period comets.

The data on actual comets are to be found in Table F and Fig. 5 of Sekanina's catalogue (1966a). These form the histogram in Fig. 8 drawn with a dashed line. The over-all agreement is good; indeed the chi-square test shows that such agreement would be expected in about half the cases if it were possible to make repetitions of this distribution.

The data on actual comets are to be found in Table F and Fig. 5 of Sekanina's catalogue (1966a). These form the histogram in Fig. 8 drawn with a dashed line. The over-all agreement is good; indeed the chi-square test shows that such agreement would be expected in about half the cases if it were possible to make repetitions of this distribution.

C. Comparison with Data: $u_a$ and $u_b$ Distributions

As noted already, these distributions are not centered on zero, and it is therefore important to include the effects of all the planets. Eight $U_a$ distributions (such as Fig. 7) must be computed. For each planet the shape of the $U_a$ distributions is different, because a given $q$ distribution (using $Q = q/a_p$) results in a different $Q$ distribution in each case. The chosen $q$ distribution is that of Table IV. Some 9000 random orbits are computed in forming these $U_a$ distributions (2000 each for Jupiter and Saturn, 1000 each for Venus, Earth, Uranus, and Neptune, and 500 each for Mercury and Mars). To each of these is applied the appropriate scale factor $(m_p/a_p)$ to convert to an analogous quantity $u_a$ measured in reciprocal a.u. A simple Monte Carlo procedure, which samples each distribution 6000 times, combines these into a single distribution. The result is the histogram drawn as a solid line in Fig. 9.

It is difficult to speak meaningfully of the standard deviation and other moments of the computed distribution. The reason is a long $1/|U|^2$ tail on both the left and the right caused by moderately close encounters of comets with planets. Indeed, of the 6000 values forming the histogram there are 60 too far on the left and 41 too far on the right to be shown on the figure, and some 28 would still be off the graph if the abscissa were extended another 1000 units of $10^{-6}$ a.u.-1 each way. Including or not including these tails can make a difference of a factor of 4 in the standard deviation. If one, quite arbitrarily, cuts off these tails at the boundaries of Fig. 9, then the average is here found to be $418 \times 10^{-6}$ a.u.-1. A theoretical average obtained by Fabry (1894) is $449 \times 10^{-6}$ a.u.-1, and later determinations (all close to these values) are discussed by Sekanina (1966a) in Tables 12 and 13 of his catalogue. Of more interest is the width and asymmetry of the distribution. With the tails cut off as described, the standard deviation $\sigma$ and the dimensionless moment coefficient of skewness $s$ [defined by $(1/N)\sum (u_a - \bar{u}_a)^3/\sigma^3$] are calcu-
lated. These values for the computed distribution are the first entry in Table V.

The histogram traced by the dashed line in Fig. 9 is the \( u_b \) distribution for 70 comets in Sekanina's catalogue (1966a). The second entry in Table V gives the moments of this \( u_b \) distribution and there is fairly good agreement with the three computed values. Also shown in Fig. 9, using the dotted line, is the negative of the \( u_b \) histogram for 81 comets in the catalogue. (All three histograms are normalized to the same area.) The third entry in Table V gives the same moments for this (negative) \( u_b \) distribution. Here the values do not agree so well, in fact the skewness agrees not at all.

The chi-square test compares the computed histogram with that for \( u_b \) in Fig. 9. The fit is reasonably good, and the test concludes that such a fit would be expected in perhaps 50% of the cases if it were possible to repeat the comparison many times using new sets of comets. However, another chi-square test comparing the computed values with the \( u_b \) values indicates no agreement whatsoever. Not one time in a thousand would such a poor fit to the computed histogram be expected on statistical grounds.

Except for the minus sign, the \( u_a \) and \( u_b \) distributions in theory should be the same (within statistical fluctuations), but this is seen not to be the case for the data available. Despite its disagreement with the \( u_b \) data, the author sees no reason to doubt the theory. The \( u_b \) distribution for the 81 comets apparently has too many large (negative) values. This is the cause of the catalogued excess of \( u \) values (where \( u = u_a + u_b \)) already mentioned in Sec. IIIA. The probable explanation is the selection effect suggested by Lyttleton and Hammersley (1963). They point out that those comets which have gained an above-average amount of total energy before perihelion (and therefore have hyperbolic osculating orbits and large negative \( u_b \) values) are just those most likely to be selected for detailed study because of the interest in possible hyperbolic original orbits. Comets so selected would not have a typical distribution of \( u_b \) values, but the \( u_a \) values for the same comets would be more representative.

### D. Correlation of \( U_a \) and \( U_b \)

It is interesting to inquire whether the value of \( U_a \) is statistically correlated with the value of \(-U_b\) for the same orbit. A scatter diagram of individual \( U_a \) and \(-U_b\) values, compared orbit by orbit for thousands of hypothetical random comets does show a kind of correlation in some cases.

For very small \( Q \) values there is a kind of "constructive interference." The trajectories are such that a comet which passes fairly near the planet on its inbound path is also likely to pass fairly near the planet on its outbound path: Apparently the relative paths are such that if it loses total energy on the inbound crossing it is also likely to lose total energy on the outbound crossing, and vice versa. When this happens, the regions on the scatter diagram are more heavily populated where \( U_a \) and \( U_b \) are both positive, or both negative. A consequence is that the histograms for \( Q=0.002 \) and 0.005 do not fit the empirical function as well as do the other cases. These two distributions are almost twice as high at moderately large \( U \) values than one would otherwise expect. The upturn near \( Q=0 \) in Fig. 6 illustrates the effect.

For \( Q=0.500 \) and extending up at least as far as \( Q=0.850 \) the correlation study indicates that a comet which loses more total energy than average before perihelion is also likely to gain more total energy than average after perihelion. This partial cancellation causes the height of the distributions at moderately large \( U \) values to be somewhat less than they would otherwise be. Despite these effects, the author could find no over-all correlation effects that would apply to random orbits where there is a wide distribution of \( Q \) values.

### Appendix A: Discussion of Tables II and III

Several comments and precautions in the use of Tables II and III should be mentioned here.

1. In several places, there are irregularities in the progression of coefficients from one entry to the next. It would be possible to smooth these, but in every case where this was tried the goodness-of-fit, as measured by least-square criteria, was worsened. The decision is to give the best fit to each distribution regardless of these jumps in the coefficients. There is the possibility that some of these can represent fairly sudden changes in the shape of the distribution and are not merely statistical fluctuations.

2. In an interpolation to find the coefficients for cases where \( Q \) and \( t \) are between those tabulated, it is best not to interpolate \( G \), but preferable to recompute...
it using Eq. (13) and the new values of $C$, $D$, $F$, and $B$. This prevents loss of exact normalization.

3. When $Q>1$, the sharp cutoff at $U_m$ causes complications in preserving normalization. Equation (13) is replaced by the more complicated expression given below in Eq. (B1), and the cases where $G=0$ then requires that $F$ be normalized using Eq. (B2) below.

**APPENDIX B: CURVE FITTING**

To fit the empirical parameters $C$, $D$, and $B$ it is best not to work with the $h$ histogram of Figs. 1 and 2 which is rather irregular, but to use instead an integrated histogram $g$, which is much smoother. Let $g(U) = N(U)/N_T$, where $N(U)$ is the number of cases for which $U$ is equal to or greater than a given $U$ value, and where $N_T$ is the total number in the sample. An equivalent analytic expression for $g$ is readily obtained by integrating Eq. (12), since $g(U) = \int_{U_m}^{U} h(U') dU'$. This gives a more complicated analytic expression for $g$ and the normalization equation is also more complicated. In particular, Eq. (13) is replaced by

$$G = 2C \left[ 1 - \frac{(F/B^2) U_m^2}{(U_m^2 + B^2)^2} \right] \frac{[\pi(\text{erfc}(\sqrt{D}) - \text{erfc}(C U_m^2 - D))]}{[\pi(\text{erfc}(\sqrt{D}) - \text{erfc}(C U_m^2 - D))]}) \right]$$

when $Q>1$. Finally, for those cases where $G=0$, the distribution depends only on $F$, $B$, and $U_m$. Then only the parameter $B$ may be adjusted, since $F$ is determined by the normalization requirement that

$$F = B^2 (B^2 + U_m^2) / U_m^2.$$  \hspace{1cm} (B2)

**REFERENCES**


---, 1967b, *ibid.* 72, 1002. Referred to as Paper II. The present work is Paper III, and another (herein called Paper IV) is to appear in *Astron. J.*


